AQUARIUS/SAC-D

Aquarius Validation Data Segment (AVDS) to Aquarius Data Processing Segment (ADPS) Interface Control Document

Project Document: **AQ-332-0331**
JPL Document: **D-35017**
18 December 2006
Initial Release

Note: Printed copies of this document may not be relied upon for official purposes. The official released version is in the Product Data Management System http://pdms.jpl.nasa.gov .

Aquarius Validation Data Segment (AVDS) to Aquarius Data Processing Segment (ADPS) Interface Control Document

PREPARED BY (CUSTODIAN):

_____ _____
P. Jeremy Werdell, (NASA GSFC) Date

_____ _____
John Gunn (AVDS manager - ESR) Date

APPROVED BY:

_____ _____
Gene C. Feldman, (Manager, Aquarius Ground System - NASA GSFC) Date

_____ _____
Gary Lagerloef, (Project Scientist, Aquarius - ESR) Date

_____ _____
David Durham, (Systems Engineer, Aquarius - NASA JPL) Date

_____ _____
Amit Sen, (Project Manager, Aquarius - NASA JPL) Date

Release History of
DOC record AQ-332-0331 revision -

View Record History | Format for Print | CLOSE

Level	Check	Sign-Off Date	Signature	Notes
OBSOLETE: No longer in Use				
RELEASED_A: Active Release				
RELEASED: Official Release	Promote Demote Comments?	09-APR-2007 11:17:04	FETTES, SANDRA G.	All required approval signatures have been obtained
(YES)	**Final CM Check**	09-APR-2007 11:16:46	FETTES, SANDRA G.	
(YES)	Document Approval	19-DEC-2006 16:49:04	FETTES, SANDRA G.	Jeremy Werdell approves document via email; see attached file
(YES)	Document Approval	19-DEC-2006 16:49:33	FETTES, SANDRA G.	John Gunn approves document via email; see attached file
(YES)	Document Approval	06-APR-2007 16:35:56	SEN, AMIT	
(YES)	Document Approval	16-FEB-2007 18:45:14	DURHAM, DAVID M.	
(YES)	Document Approval	11-JAN-2007 06:03:31	FELDMAN, GENE C.	
(YES)	Document Approval	10-JAN-2007 14:16:53	LAGERLOEF, GARY S.	

DOCUMENT CHANGE LOG

Change Number	Change Date	Pages Affected	Changes/ Notes	General Comments
Draft	7/25/2006	All	Preliminary Release	First complete draft prepared
Draft	8/15/2006	All	Text added to reflect comments from July 2006 pre-CDR meeting.	
Draft	10/6/2006	Page 4	Text added to reflect uncertainty in format for name of data output file	
-	12/18/2006	All	Formatted for Document template	Initial Release

Table of Contents

1. Introduction

1.1 Purpose and Scope

The Aquarius Validation Data Segment (AVDS; Earth & Space Research, Seattle, WA) will statistically evaluate the Aquarius sea surface salinity (SSS) data product using a global set of *in situ* salinity measurements. These *in situ* data are collected regularly by a number of ongoing field measurement programs (external to the Aquarius program) conducted over the breadth of the world's oceans. This activity will ensure that the Aquarius instrument and data processing system yield satellite-derived SSS observations that meet the mission Science Requirements and are comparable and consistent with other marine observation programs.

This document describes the responsibilities of both the AVDS and the Aquarius Data Processing Segment (ADPS; NASA Goddard Space Flight Center, Greenbelt, MD) in a data exchange required as part of this validation effort. The scenario described herein is based on the current understanding of the technologies involved and the interfaces required to accomplish this task. If mutually agreed upon alternative methods are deemed superior, both the AVDS and ADPS will work towards implementing those alternatives.

The rationale for the data exchange is to streamline the AVDS validation activity by using the preexisting ADPS hardware, software, and satellite data processing facilities (and thereby eliminating a redundant effort if the AVDS were to develop parallel facilities).

1.2 Background

The Aquarius/SAC-D Mission is designed to provide monthly, global 150-km resolution, SSS maps. Briefly, the Aquarius instrument measures the brightness temperature (T_B) of the sea surface at L-band, 1413 MHz. These data, in combination with coincident sea surface temperatures, radar scatterometer measurements, and other ancillary data, are used to estimate SSS (Lagerloef et al., 1995). The ancillary data are acquired from sources other than the Aquarius instrument and are used to correct or geo-locate instrument measurements. The Aquarius instrument will be hosted with a number of other science instruments on the SAC-D service platform, built by CONAE, the Argentine government's space agency.

2. AVDS Responsibilities

The AVDS will:

1. acquire *in situ* SSS data (Section 2.1)
2. evaluate the *in situ* SSS data (Section 2.2)

3. reformat the *in situ* SSS data (Section 2.3)
4. deliver the *in situ* SSS data to the ADPS (Sections 2.3, 2.4 and 3.1) and to the Aquarius Science Team (AST) and general oceanographic community
5. acquire coincident Level-2 Aquarius data from the ADPS (Sections 2.5 and 3.3)
6. statistically evaluate, filter and bin as necessary, tabulate, and redistribute optimally matched Level-2 Aquarius and *in situ* SSS data to the AST and general oceanographic community

2.1 *In situ* data acquisition

The AVDS will assemble *in situ* SSS data as it becomes available from various sources. A number of these sources have been identified, but others will be added as new data become available and additional marine observation programs reach maturity. The data sources will be described in detail in the AVDS Development Plan (AVDS-DP). In general, the marine observation programs are international in nature and the data are often publicly released in near-real time. Since each external data source makes use of a unique data file format, the AVDS will consolidate these data into the consistent data file format described in Section 2.3.

Data from all available *in situ* data sources will be acquired daily.

2.2 *In situ* data evaluation

The AVDS will apply quality control metrics to the *in situ* data as described in the AVDS-DP.

2.3 *In situ* data file format

The AVDS will prepare *in situ* data files following standard SeaWiFS Bio-optical Archive and Storage System (SeaBASS; http://seabass.gsfc.nasa.gov) protocols. This will ensure the efficient and seamless integration of these data into the existing ADPS validation infrastructure and permit the use of existing ADPS analysis software.

A complete description of the SeaBASS data file format is available in Werdell and Bailey (2002). Briefly, files are flat ASCII text to ensure compatibility with multiple operating systems and processing languages. Each file consists of a matrix of data values preceded by a descriptive series of predefined metadata headers (Figure 1). Each header consists of a paired keyword and value, one per line, in the form `/<keyword>=<value>` (see Appendix A). Note the slash (`/`) that precedes each keyword.

At a minimum, the ADPS requires the data matrix to include the following set of columns: year, month, day, hour, minute, second, latitude (decimal degrees), longitude (decimal degrees), delta time (in seconds), delta space (search radius in kilometers), and a record locator code or station / buoy reference code. The latter is required to facilitate communication between

the ADPS and AVDS if problematic data points need to be identified. The date and time fields are UTC. The delta values define temporal and spatial windows for selecting acceptable T_B and derived SSS data from the satellite. Additional columns may include SSS and ancillary data products, such as sea surface temperature and wind speed, if desired. As this process matures, additional ancillary data will be provided to assist with subsequent analysis of the satellite-to-*in situ* validation results. The `/fields` and `/units` headers identify each column and have been standardized (see Appendix B) to facilitate integration with the existing ADPS match-up software.

A single file will contain all *in situ* data specifications for a given day. The files will be named following 'AVDSYYYYDDDHHMMSS.txt', where YYYY is the year, DDD the day of the year, HH the hour, MM the minute, and SS the second (all in UTC) the file is created.

2.4 Delivery of *in situ* data

New *in situ* data files will be made available daily for download from an AVDS File Transfer Protocol (FTP) server at <ftp://AVDS.esr.org> [TBC] for retrieval by the ADPS. New files will be immediately identifiable using the file naming conventions described in Section 2.3. Given the extent of the quality control metrics described in Section 2.2, an acquisition-to-upload lag time of [TBD] is expected.

2.5 Satellite data acquisition

The AVDS will non-interactively download extracted Aquarius data files and related ancillary data products (e.g., using 'cron' and 'ncftp') from the ADPS FTP server (see Section 3.3) on a daily basis. These satellite files will in turn be used with the *in situ* observations to statistically evaluate the Aquarius SSS data product (see the AVDS-DP).

3. ADPS Responsibilities

The ADPS will:

1. acquire *in situ* SSS data from the AVDS (Section 3.1 and 2.4)
2. identify coincident Aquarius observations using the *in situ* targets (Section 3.2)
3. distribute extracted Aquarius SSS data files and related ancillary data products to the AVDS (Section 3.3 and 2.5)

3.1 *In situ* data acquisition

The ADPS will non-interactively download *in situ* SSS data files (e.g., using 'cron' and 'ncftp') from the AVDS FTP server (see Section 2.4) on a daily basis.

3.2 Identification and extraction of coincident satellite data

The ADPS will use their existing validation infrastructure to identify coincident satellite and *in situ* SSS observations (also referred to as 'match-ups'). A complete description of this process is available in Bailey and Werdell (2006). Briefly, the infrastructure consists of three tiers: physical data archives on which the satellite and *in situ* data files reside; a relational database management system (RDBMS; currently the SQL Server product from Sybase, Inc.) to catalog and store relevant metadata from these files; and a collection of front-end control software, written in Practical Extraction and Report Language (PERL) and Transact Structured Query Language (SQL). The system currently operates on a LINUX platform (CentOS 4.2).

In general, the front-end software will be used to customize calls to the RDBMS and retrieve the satellite data files. The RDBMS will be used to locate coincident satellite files, given a list of *in situ* metadata and user-defined temporal and spatial thresholds (the delta values described in Section 2.3). Once the *in situ* and satellite observation times and locations are registered, RDBMS queries will allow both the location of satellite measurements within specified space and time windows of the *in situ* observation, and the location of *in situ* measurements within specified space and time windows of the satellite observation (reverse search). The satellite files will be copied locally, extracted to a reduced geographic region specified by the search criteria, and processed to Level-2 following mutually agreed upon acceptance requirements and satellite data processing protocols (see ADPS Report TBD). The geographic extent of the extract files (in number of pixels) is to be dependent on the spatial-temporal search radius provided by the AVDS. The extract files will include Aquarius-derived Level 2 data, related metadata (e.g., sensor and solar zenith angles), and coincident ancillary data (e.g., sea surface temperature and wind speed), as mutually agreed upon by the AVDS and ADPS. Level 2 data includes both the Level 2A geolocated, calibrated sensor units at the surface (0.72 sec sample interval), and the Level 2B geolocated SSS derived from Level-2A values (averaged to a 5.76 sec sample interval), as specified in the Aquarius Science Algorithm Requirements Document.

The ADPS will retain a similar file naming convention to that of the *in situ* data (i.e., 'ADPSYYYYDDDHHMMSS_XXXX.hdf') to maintain continuity with these data and to make new data files readily identifiable. As multiple satellite files will be associated with each *in situ* data file (up to one per field data record provided in the file), the XXXX in the satellite file name will indicate the record locator code or station / buoy reference code described in Section 2.3. For the case of the reverse search this file name convention may be ambiguous so the file name for that search output is TBD.

The ADPS has the additional capability to provide Level-1 data should that information be deemed requisite in the future.

3.3 Distribution of satellite data

New satellite data files will be made available daily for download from an ADPS FTP server at <ftp://oceans.gsfc.nasa.gov> [TBC] for retrieval by the AVDS. New files will be immediately identifiable using the file naming conventions described in Section 3.2. Given the processing time required to accomplish that described in Section 3.2, an satellite data processing-to-upload lag time of [TBD] is expected.

4. Operational Scenario

An operational scenario is described below (see also Figure 2).

1. The AVDS acquires *in situ* SSS data from an independent marine observation program.
2. The AVDS evaluates and reformats these *in situ* SSS data.
3. The AVDS uploads these *in situ* data to their FTP server.
4. The ADPS downloads these *in situ* data from the AVDS FTP server.
5. The ADPS uses these *in situ* data to identify coincident Aquarius satellite data files.
6. The ADPS extracts the Level-1 satellite data from a reduced, geographic area centered on the *in situ* station.
7. The ADPS processes to Level-2 the extracted, coincident satellite data.
8. The ADPS adds relevant metadata and ancillary data to the extracted Level-2 files.
9. The ADPS uploads the extracted Level-2 satellite data to their FTP server.
10. The AVDS downloads the satellite data from the ADPS FTP server.
11. The AVDS provides the coincident satellite and *in situ* data to the AST to statistically evaluate the Aquarius SSS data product.

5. Development, Simulation, and Schedule

5.1 Development

The development of this data exchange will be incremental and dependent on the completion of other Aquarius system tasks. For example, completion of the satellite-to-*in situ* match-up software relies upon the final definition of Level-2 data format. Development of the infrastructure, however, can move forward in parallel until requisite data formats, and other software dependencies, are finalized.

5.2 Simulation

Modeled data will be used to test the match-up software and to streamline the coordination of the data exchange between the AVDS and the ADPS prior to the satellite launch date. In this exercise, Level-2 data products will be generated using simulated Level-1 data [responsibilities of the AVDS and ADPS TBD]. This simulation will be conducted as part of a large-scale simulation of the entire Aquarius processing scheme. The intent of the latter is to use a synthesized data stream to run a system-wide simulation for one-year prior to launch to test all system components. Testing of the AVDS / ADPS interaction described herein will begin approximately 6 months prior to the initialization of the system-wide simulation (or, about 18 months before launch). [We should outline

5.3 Preliminary schedule

34 months before launch (bl) (~Sep 2006)	- ICD agreed upon
28 months bl (~Mar 2007)	- data formats finalized
18 months bl (~Jan 2008)	- AVDS/ADPS data exchange testing begins
12 months bl (~Jul 2008)	- system-wide simulation begins
0 months bl (~Jul 2009)	- satellite launch

6. References

Bailey, S.W. and P.J. Werdell (2006): A multi-sensor approach for the on-orbit validation of ocean color satellite data products. *Remote Sensing of Environment*, **102**, 12-23.

Lagerloef, G., C. Swift, and D. LeVine (1995): Sea surface salinity: The next remote sensing challenge. *Oceanography*, **8**, 44-50.

Werdell, P.J. and S.W. Bailey (2002): The SeaWiFS Bio-optical Archive and Storage System (SeaBASS): Current Architecture and Implementation. *NASA Technical Memorandum 2002-211617*, NASA Goddard Space Flight Center, Greenbelt, Maryland, pp. 45.

7. Figures

Figure 1. Graphical representation of the structure of a SeaBASS data file.

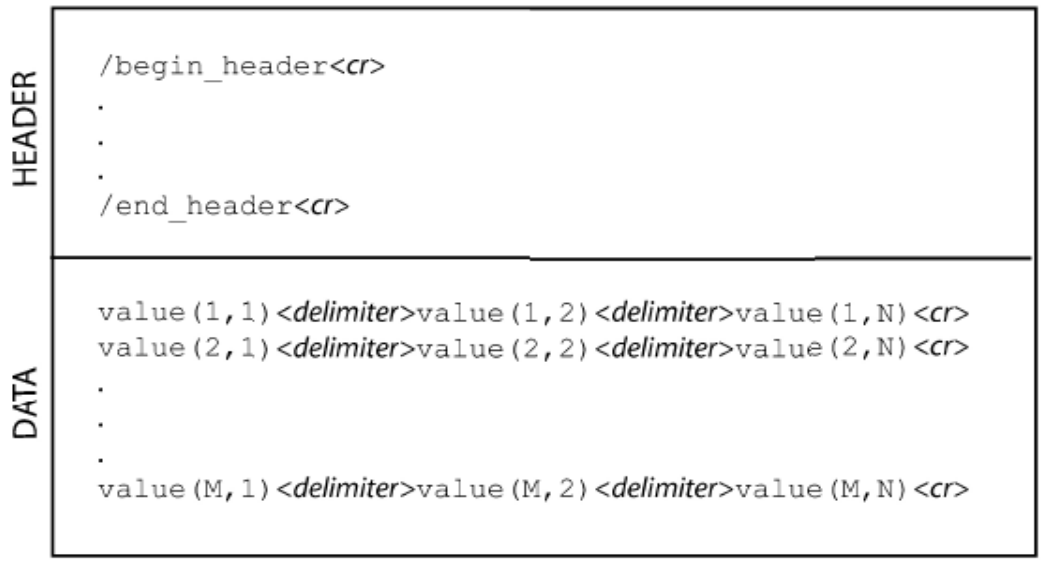
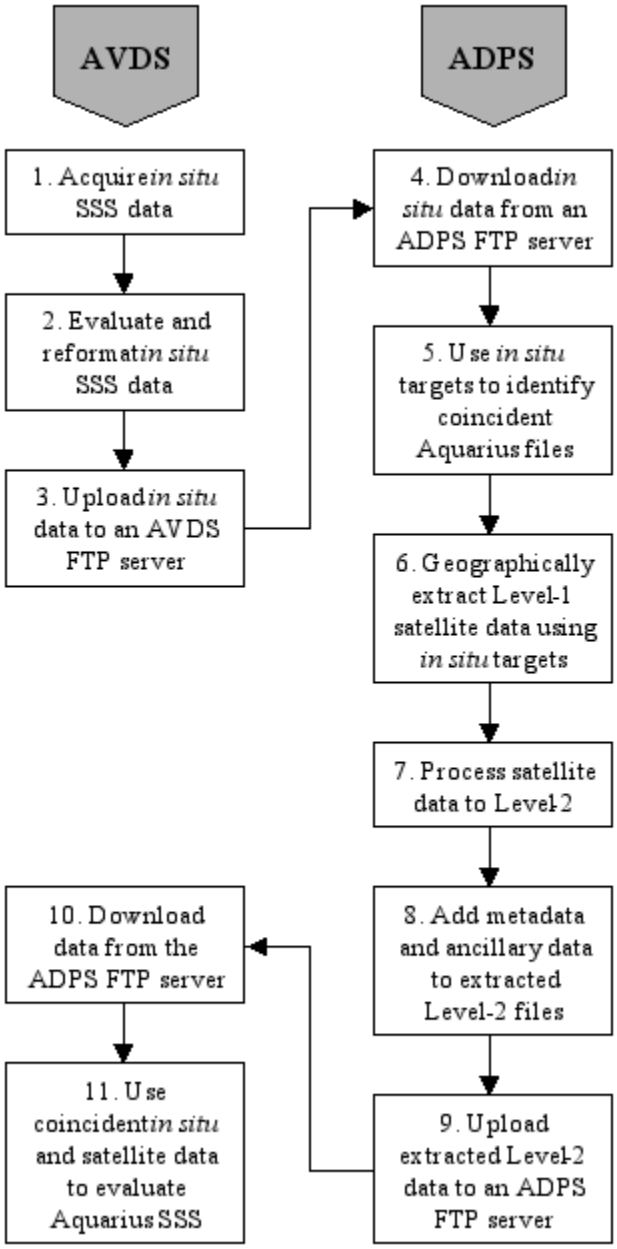


Figure 2. Graphical representation of Section 4. The left and right columns indicate the AVDS and ADPS responsibilities, respectively. The numbers in each box correspond to the numbered list in Section 4.



8. Appendices

8.1 Appendix A

The AVDS / ADPS data exchange requires a reduced set of SeaBASS metadata headers. The header keywords and a description of each are provided in the table below. Several headers in this reduced set are not required. For a full list of SeaBASS headers, please refer to Werdell and Bailey (2002).

HEADER	REQUIRED	DESCRIPTION
`/begin_header`	Y	The first line of every data file, indicating the beginning of the header block. This header does not have an input argument.
`/experiment`	N	The name of the long-term research project, e.g., TOGA and PIRATA, if applicable.
`/cruise`	N	The name of the specific cruise, or subset of the experiment, where the data in the file were collected, if applicable.
`/station`	N	The name of the station or deployment area where data in the file were collected, if applicable.
`/documents`	N	A list of accompanying documentation, if applicable.
`!`	N	A space for additional comments. Common comments include additional ancillary information about the data file, sea and sky states, difficulties encountered during data collection, methods of data collection, instruments used, and a description of nonstandard SeaBASS field names included in the data file.
`/missing`	Y	The null value used as a numeric placeholder for any missing data in the data file. Each row of data must contain the same number of columns as defined in the `/fields` and `/units` headers. Only one missing value is allowed per file. It is required that this value be non-zero.
`/delimiter`	Y	The delimiter of the columns of data. Accepted delimiters include tab, space, and comma. Only a single delimiter is permitted per data file.
`/fields`	Y	A list of the fields, e.g., `lat`, for each column of data included in the data file. Each entry describes the data in a single column, and every column must have an entry.
`/units`	Y	A list of the units, e.g., `degreesC`, for each column of data included in the data file. Every value in `/fields` must have an appropriate value listed in this header.
`/end_header`	Y	The final line of the header block, indicating the beginning of the data block. This header does not have an input argument.

8.2 Appendix B

The SeaBASS data format requires the use of series of standardized field names and units for the `/fields` and `/units` headers, respectively. These are used to uniquely, and unequivocally, identify each column in the data matrix. Non-standard names and units are permissible when and standardized version does not exist. For a full list of SeaBASS standardized field names and units, please refer to Werdell and Bailey (2002).

FIELD	UNIT	DESCRIPTION
day	dd	Day of month (1 to 31).
hour	hh	Hour (0 to 23), in UTC.
id	none	Record locator or identification code.
lat	degrees	Latitude, in decimal degrees. Locations south of the equator have negative values
lon	degrees	Longitude, in decimal degrees. Locations west of the Prime Meridian have negative values.
minute	mn	Minute (0 to 59), in UTC.
month	mo	Month (1 to 12).
second	ss	Second (0 to 59), in UTC.
sss	psu	Sea surface salinity.
sst	degreesC	Sea surface temperature.
wind_speed	m/s	Wind speed.
year	yyyy	Year, in four digits.